Acta Cryst. (1973). B**29**, 1234

New Investigations of Cytosine and Its Monohydrate

BY R. J. MCCLURE AND B. M. CRAVEN

Department of Crystallography, University of Pittsburgh, Pittsburgh, Pa 15260, *U.S.A.*

(*Received* 20 *November* 1972; *accepted* 30 *January* 1973)

Cytosine is orthorhombic, space group $P2_12_12_1$, with $a=13{\cdot}044$ (2), $b=9{\cdot}496$ (1), $c=3{\cdot}814$ (1) Å, $Z=4$. Cytosine monohydrate is monoclinic, space group $P2_1/c$, with $a=7{\cdot}783$ (2), $b=9{\cdot}825$ (2), $c=7{\cdot}668$ (2) Å, $\beta=99°34$ (1)′, $Z=4$. The bond lengths of the carbonyl group, previously reported to be different in cytosine and the monohydrate, are presently found to be 1·241 and 1·251 Å, respectively. This and other differences in molecular structure are of marginal significance.

Introduction

The crystal structures of cytosine (Barker & Marsh, 1964; hereafter BM) and its monohydrate (Jeffrey & Kinoshita, 1963; hereafter JK) have been determined from photographic intensity data. It was found that the carbonyl bond length was significantly longer (1·260 *vs.* 1·234 Å) in the monohydrate. This difference was attributed to the effect of hydrogen bonding. Similar but smaller (0·01 Å) effects have been reported in a series of barbiturate crystal structures (Craven, Cusatis, Gartland & Vizzini, 1973). Barbiturate bond length differences are associated with greater hydrogen bonding differences than are found in cytosine and its monohydrate. We have redetermined the latter structures in an attempt to resolve this inconsistency.

Experimental

Crystals of cytosine and its monohydrate were obtained as described by BM and JK. The crystal of cytosine which was selected for data collection measured $0{\cdot}3\times0{\cdot}3\times0{\cdot}6$ mm and was mounted in an arbitrarily chosen orientation, with the longest crystal dimension (c^*) approximately 20° from the diffractometer φ axis. The monohydrate crystal measured $0{\cdot}2\times0{\cdot}2\times0{\cdot}5$ mm and was mounted with the longest dimension (b^*) approximately 19° from the diffractometer φ axis. The X-ray data were measured using a computer controlled four-circle diffractometer with nickel-filtered Cu $K\alpha$ radiation ($\lambda=1{\cdot}5418$ Å). The lattice parameters are in good agreement with those of BM but are smaller than those of JK by about 0·02 Å, possibly because the latter did not take film shrinkage into account.

Intensity data were collected in the asymmetric region of reciprocal space with $2\theta_{max}=130°$ using $\theta/2\theta$ scans at 30 sec deg^{-1} 2θ and with background counts of 10 sec at each scan limit. In the case of the monohydrate, the use of a smaller crystal gave a higher proportion of unobservably weak reflections. For this reason, a second data set was measured, using the same scan rate but with 20 sec background counts. The two sets were averaged with equal weights for integrated intensities greater than 1·5 $\sigma(I)$. Data collection was monitored by the use of three standard reflections which exhibited no significant intensity changes. No corrections were made for X-ray absorption or extinction. Integrated intensities less than 1·5 $\sigma(I)$ were considered to be unobservably weak and were assigned values of $\sigma(I)$.

The atomic parameters of BM and JK were refined by a full-matrix least-squares procedure. The function $\sum w(\Delta F)^2$ was minimized, using the weights given in

Table 1. *Intensity data and refinement criteria*

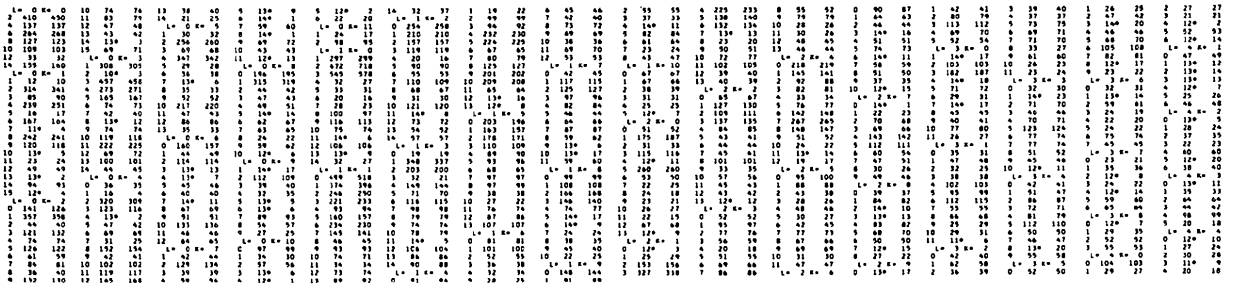
	Anhydrous	Monohydrate
Number of intensity data:		
(*a*) Observed	431	749
(*b*) Unobservably weak	74	129
Weighting scheme, $w(F)=1/\sigma^2(F)$		
$\sigma^2(F)=A+BF+CF^2$		
A	0·231	0·185
B	−0·535	−0·400
C	0·0036	0·00325
Standard deviation of observation of unit weight	1·18	1·03
R index,* including unobserved reflections	0·039	0·047
R index, excluding unobserved reflections	0·031	0·037
Weighted R index, including unobserved reflections	0·040	0·051
Weighted R index, excluding unobserved reflections	0·034	0·045

* $R=\sum_h |\Delta F|/\sum_h |F_{meas}|$ where $\Delta F=|F_{meas}|-|F_{calc}|$.

Table 2. *Observed and calculated structure factors*

Successive columns contain h, $10|F_{obs}|$ and $10F_{calc}$. Reflections marked with an asterisk were unobserwably weak.

(*a*) Anhydrous cytosine



(*b*) Cytosine monohydrate. Reflections marked + were excluded from the final refinement procedure.

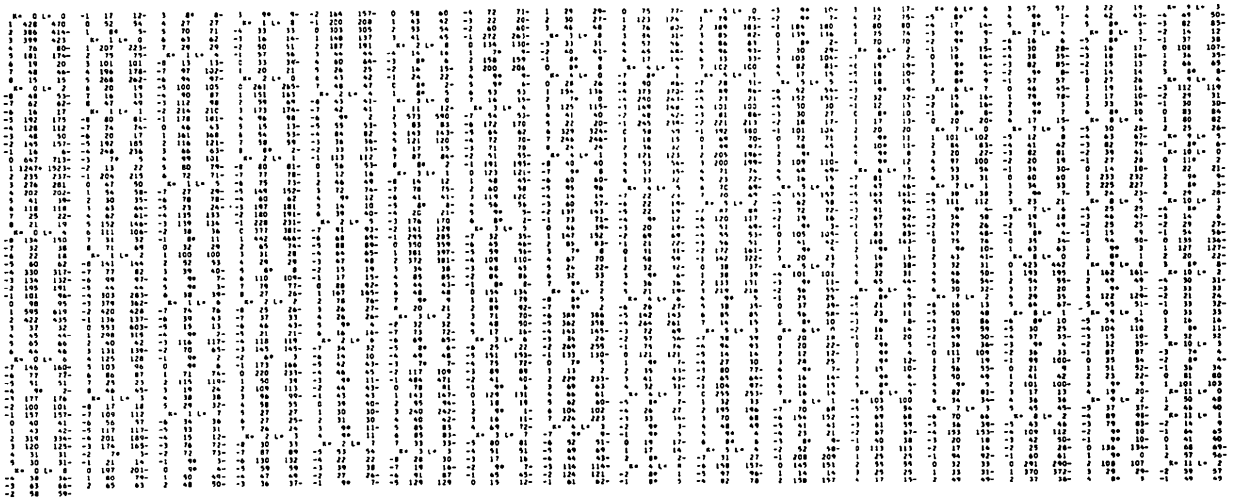


Table 3. *Atomic parameters*

The parameters are the fractional atomic coordinates and anisotropic temperature factors. The latter are defined by $T=\exp[-\sum_i\sum_j h_i h_j \beta_{ij}]$. The e.s.d.'s, which are given in parentheses, refer to the least significant digit in the corresponding parameter value. The parameters for anhydrous cytosine are above those for cytosine monohydrate.

(*a*) Heavier atoms

	10^4x	10^4y	10^4z	β_{11}	β_{22}	β_{33}	β_{12}	β_{13}	β_{23}
N(1)	222 (1)	285 (2)	4410 (6)	0·0033 (1)	0·0054 (2)	0·053 (2)	−0·0001 (1)	−0·0002 (4)	−0·0005 (5)
	843 (2)	76 (1)	2172 (2)	0·0111 (3)	0·0044 (2)	0·0153 (3)	−0·0002 (1)	0·0046 (2)	0·0001 (2)
C(2)	−164 (2)	1561 (2)	3276 (7)	0·0029 (1)	0·0052 (2)	0·044 (2)	0·0002 (1)	0·0010 (4)	0·0001 (6)
	016 (2)	1281 (1)	2391 (2)	0·0101 (3)	0·0053 (2)	0·0121 (3)	0·0005 (2)	0·0035 (2)	−0·0003 (2)
O(2)	−998 (1)	1595 (2)	1710 (5)	0·0028 (1)	0·0060 (2)	0·066 (2)	−0·0004 (1)	−0·0024 (3)	0·0002 (5)
	−1466 (1)	1239 (1)	2821 (2)	0·0124 (2)	0·0070 (1)	0·0209 (3)	0·0000 (1)	0·0079 (2)	−0·0001 (1)
N(3)	402 (1)	2745 (2)	3877 (5)	0·0028 (1)	0·0048 (2)	0·051 (2)	0·0000 (1)	−0·0002 (4)	0·0007 (5)
	795 (2)	2470 (1)	2108 (2)	0·0101 (3)	0·0047 (1)	0·0143 (3)	0·0006 (1)	0·0042 (2)	0·0003 (1)
C(4)	1308 (2)	2635 (3)	5492 (7)	0·0030 (1)	0·0059 (3)	0·040 (3)	−0·0003 (1)	0·0007 (4)	−0·0003 (6)
	2348 (2)	2445 (2)	1570 (2)	0·0104 (3)	0·0057 (2)	0·0118 (3)	−0·0001 (2)	0·0029 (2)	0·0003 (2)
N(4)	1842 (2)	3828 (2)	5967 (7)	0·0034 (1)	0·0069 (3)	0·077 (2)	−0·0005 (1)	−0·0035 (5)	−0·0004 (7)
	3097 (2)	3625 (1)	1312 (2)	0·0121 (3)	0·0051 (2)	0·0218 (2)	−0·0003 (2)	0·0079 (3)	0·0003 (2)
C(5)	1705 (2)	1337 (2)	6674 (8)	0·0031 (1)	0·0075 (3)	0·046 (2)	0·0008 (2)	−0·0014 (5)	−0·0004 (7)
	3188 (2)	1198 (2)	1270 (2)	0·0108 (3)	0·0065 (2)	0·0163 (3)	0·0012 (2)	0·0062 (3)	−0·0004 (2)
C(6)	1134 (2)	187 (2)	6084 (7)	0·0038 (1)	0·0054 (3)	0·051 (2)	0·0012 (1)	0·0015 (5)	0·0019 (6)
	2399 (2)	45 (2)	1603 (2)	0·0123 (3)	0·0053 (2)	0·0160 (4)	0·0015 (2)	0·0044 (3)	−0·0005 (2)
O(W)	—	—	—	—	—	—	—	—	—
	−3534 (2)	2929 (1)	260 (2)	0·0129 (2)	0·0099 (2)	0·0201 (3)	0·0004 (2)	0·0067 (2)	0·0015 (2)

Table 3 (*cont.*)

(*b*) Hydrogen atoms

	x	*y*	*z*
H(1)	−0·013 (2)	−0·057 (3)	0·363 (7)
	0·031 (2)	−0·067 (2)	0·236 (2)
H(3)	0·154 (2)	0·463 (3)	0·541 (9)
	0·259 (3)	0·434 (2)	0·157 (2)
H(4)	0·248 (2)	0·378 (2)	0·652 (8)
	0·408 (3)	0·361 (2)	0·098 (3)
H(5)	0·237 (2)	0·128 (2)	0·796 (8)
	0·420 (3)	0·122 (2)	0·087 (3)
H(6)	0·132 (2)	−0·078 (3)	0·692 (8)
	0·285 (2)	−0·086 (2)	0·148 (2)
H*W*(1)	—	—	—
	−0·294 (3)	0·250 (2)	0·105 (3)
H*W*(2)	—	—	—
	−0·285 (3)	0·323 (2)	−0·039 (3)

Table 1. Atomic scattering factors for carbon, nitrogen and oxygen from *International Tables for X-ray Crystallography* (1962) were used. The values of Stewart, Davidson & Simpson (1965) were used for hydrogen atoms. Thermal parameters for each hydrogen atom were assumed to be the same as the atom to which it is bonded. The refinement converged satisfactorily (Table 1).

Observed and calculated structure amplitudes are given in Table 2 and atomic parameters in Table 3. The anisotropic thermal parameters were analyzed in terms of the rigid-body model (Schomaker & Trueblood, 1968). A comparison of experimental (U_{meas}) and calculated (U_{calc}) anisotropic atomic thermal parameters shows that the assumed rigid-body vibration is a fair approximation, particularly for the anhydrous crystal structure. Thus the values of $[\sum(U_{meas}-U_{calc})^2/(n-s)]^{1/2}$ are 0·0014 Å^2 for the anhydrous and 0·0017 Å^2 for the monohydrate structures. For the anhydrous crystal structure there are 8 U_{ij} values out of 48 for which $|U_{meas}-U_{calc}|/\sigma(U_{meas})$ exceeds 2. The corresponding number is 19 for the monohydrate. The poorest agreement ($4{\cdot}2\sigma$) is in U_{22} for the atom N(4) in the monohydrate. The librational tensor components for the cytosine molecule in the two crystal structures are very similar when they are compared with respect to the same molecular axial system (Table 4). The greatest difference in the resulting bond length corrections is for C(6)–N(1), the values being 0·004 Å in anhydrous cytosine and 0·006 Å in the monohydrate. The corrections

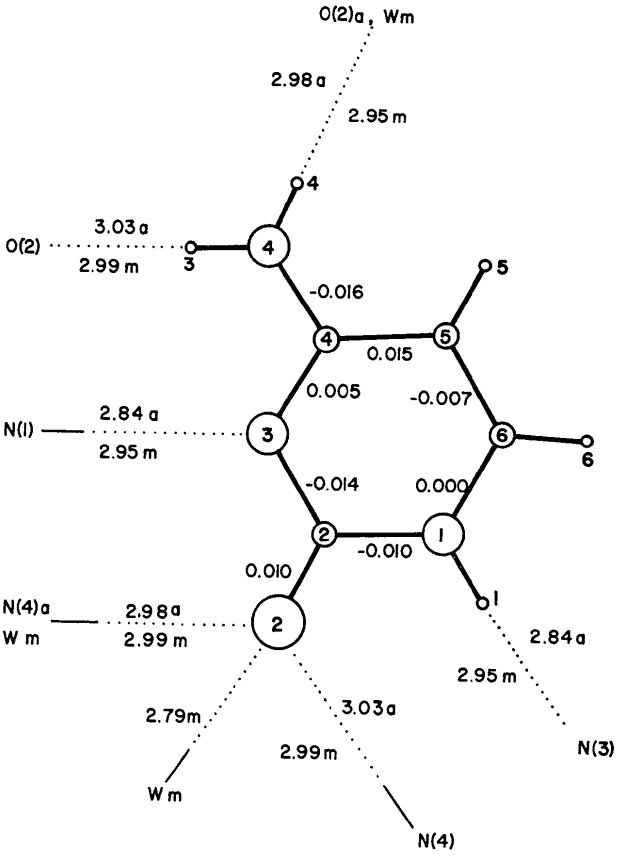


Fig. 1. Comparison of cytosine in the anhydrous and monohydrated crystal structures. Circles of decreasing size represent oxygen, nitrogen, carbon and hydrogen atoms. Atomic numbering is consistent with Table 3. Numbers alongside the C–O, C–N and C–C bonds are the bond length difference (monohydrate–anhydrous) in Å. The hydrogen bonding environment of the cytosine molecule in both structures is represented schematically. Atom labels and N···O or N···N hydrogen bond lengths are followed by **a** or **m** denoting anhydrous or monohydrate respectively. Atoms not labeled in this way are the same in both crystal structures.

Table 4. *Rigid-body thermal motion analysis*

The tensors **L** for libration, **T** for translation and **S** for screw motion are referred to an origin at the molecular centroid with fractional coordinates (0·0681, 0·1771, 0·469) and (0·1402, 0·1547, 0·1905) in the anhydrous and monohydrate structures respectively. Tensor components are with respect to an orthogonal molecular axial system with axis (1) the normal to the molecular plane, axis (2) along the vector C(6) → N(3), and axis (3) completing a right-handed system.

	Cytosine			Cytosine monohydrate		
L, σ(**L**) (degrees^2)	5·7 (1·6)	−1·5 (1·6)	−2·0 (1·8)	5·6 (1·6)	−0·2 (1·8)	1·2 (2·1)
	−1·5	12·6 (1·7)	2·4 (1·8)	−0·2	16·9 (2·1)	5·1 (2·3)
	−2·0	2·4	20·4 (2·2)	1·2	5·1	22·6 (2·9)
(L_1, L_2, L_3)	21·5	12·2	5·3	24·7	16·8	5·6
T, σ(**T**) (Å^2)	0·0238 (12)	0·0023 (8)	0·0026 (9)	0·0315 (15)	0·0005 (10)	0·0019 (11)
	0·0023	0·0237 (8)	0·0005 (8)	0·0005	0·0223 (9)	0·0008 (9)
	0·0026	0·0005	0·0265 (8)	0·0019	0·0008	0·0257 (9)
T_1, T_2, T_3	0·029	0·024	0·021	0·032	0·026	0·022
S, (degree Å)	−0·01	−0·04	0·09	0·00	0·01	0·05
	0·01	−0·04	−0·05	0·02	0·00	−0·01
	0·08	−0·03	0·05	0·03	−0·02	0·00

for the bond length C(2)–O(2) are 0·004 and 0·005 Å, respectively.

Discussion

We have chosen an atomic numbering system which is different from those of BM and JK, but which is more consistent with chemical usage (Fig. 1). Bond lengths and angles for the cytosine molecule are listed in Table 5, together with the values reported by BM and JK. Our values for the water molecule in cytosine are 0·84 and 0·82 Å for the O–H bond lengths and 106° for the H–O–H angle.

The differences in bond lengths of cytosine in the anhydrous and monohydrate crystal structures are shown in Fig. 1. The largest difference (0·016 Å), which is in the bond length C(4)–N(4) is significant (3·5σ) in terms of the e.s.d.'s derived from the least-squares refinement. However, these are probably underestimated, and it is unlikely that any of the bond length differences, including that of C(2)–O(2) (0·010 Å), are more than possibly significant.

Although the crystal structures of cytosine and its monohydrate are quite different. Fig. 1 shows that the hydrogen bonding interactions of the cytosine molecule are similar, particularly at atom N(4). The N(1)H···N(3) hydrogen bond is stronger in the anhydrous crystal structure, as shown by the shorter N(1)···N(3) distance (2·84, 2·95 Å). Of the hydrogen bonds formed at O(2), the N(4)H(3)···O distances are similar (3·03, 2·99 Å). The N(4)H(4)···O distance (2·98 Å) in the anhydrous structure and the corresponding O–H···O distance (2·79 Å) in the monohydrate are both about 0·1 Å longer than the distance corresponding to the peak in the observed distribution of hydrogen bonded N···O and O···O distances (Pimentel & McClellan, 1959; Voet & Rich, 1970). The greatest difference thus lies in the second OH–O hydrogen bond in the monohydrate. This is a very weak interaction (O···O distance 2·99 Å), but it has no counterpart in the anhydrous structure.

Thus, bond length differences which might be attributed to hydrogen bonding effects in cytosine and

Table 5 *Bond lengths and angles*

These parameters are not corrected for the effect of thermal motion.

	Cytosine		Cytosine monohydrate	
	BM	Present study	JK	Present study
	σ=0·003 Å	σ=0·003 Å	σ=0·004 Å	σ=0·002 Å
N(1)–C(2)	1·374 Å	1·381 Å	1·376 Å	1·371 Å
C(2)–N(3)	1·364	1·364	1·354	1·350
N(3)–C(4)	1·337	1·336	1·351	1·341
C(4)–C(5)	1·424	1·410	1·432	1·425
C(5)–C(6)	1·342	1·340	1·348	1·333
C(6)–N(1)	1·357	1·353	1·361	1·353
C(2)–O(2)	1·234	1·241	1·260	1·251
C(4)–N(4)	1·330	1·342	1·332	1·326
	σ=0·04 Å	σ=0·03 Å	—	σ=0·02 Å
N(1)–H(1)	0·88 Å	0·98 Å	0·87 Å	0·86 Å
N(4)–H(3)	0·87	0·89	0·89	0·85
N(4)–H(4)	0·86	0·86	0·83	0·84
C(5)–H(5)	0·87	0·99	0·85	0·89
C(6)–H(6)	1·01	1·01	0·87	0·97
	σ=0·2°	σ=0·2°	σ=0·4°	σ=0·15°
C(2)–N(1)–C(6)	122·7°	121·9°	121·3°	121·5°
N(3)–C(2)–N(1)	118·1	118·2	120·1	119·6
C(2)–N(3)–C(4)	119·9	119·4	118·9	119·1
N(3)–C(4)–C(5)	122·0	122·7	122·0	121·7
C(4)–C(5)–C(6)	117·3	117·0	117·1	117·5
C(5)–C(6)–N(1)	120·1	120·8	120·6	120·5
N(1)–C(2)–O(2)	119·8	119·5	118·4	118·5
N(3)–C(2)–O(2)	122·2	122·2	121·5	121·9
N(3)–C(4)–N(4)	118·2	117·1	117·8	118·1
C(5)–C(4)–N(4)	119·9	120·2	120·2	120·2
	σ=2°	σ=2°	—	σ=1·4°
C(2)–N(1)–H(1)	120°	117°	116°	118°
C(6)–N(1)–H(1)	117	120	123	121
C(4)–N(4)–H(3)	123	120	119	117
C(4)–N(4)–H(4)	124	117	118	118
H(3)–N(4)–H(4)	114	123	122	124
C(4)–C(5)–H(5)	123	121	126	120
C(6)–C(5)–H(5)	119	122	117	123
C(5)–C(6)–H(6)	122	124	123	125
N(1)–C(6)–H(6)	118	115	116	114

cytosine monohydrate should be of marginal significance, in agreement with our findings.

This work was supported by Grant Nos. NS-02763 and GM-01728 of the U.S. Public Health Service, National Institutes of Health. The IBM 1130 and 7090 computer programs used were those written or modified by Dr R. Shiono and J. Rutherford.

References

BARKER, D. L. & MARSH, R. E. (1964). *Acta Cryst.* **17**, 1581–1587.

CRAVEN, B. M., CUSATIS, C., GARTLAND, G. L. & VIZZINI, E. A. (1973). *J. Mol. Struct.* In the press.

International Tables for X-ray Crystallography (1962). Vol. III. Birmingham: Kynoch Press.

JEFFREY, G. A. & KINOSHITA, Y. (1963). *Acta Cryst.* **16**, 20–38.

PIMENTEL, G. C. & MCCLELLAN, A. L. (1959). *The Hydrogen Bond.* San Francisco: Freeman.

SCHOMAKER, V. & TRUEBLOOD, K. N. (1968). *Acta Cryst.* B**24**, 63–76.

STEWART, R. F., DAVIDSON, E. R. & SIMPSON, W. T. (1965). *J. Chem. Phys.* **42**, 3175–3187.

VOET, D. & RICH, A. (1970). *Prog. Nucleic Acid Res. Mol. Biol.* **10**, 183–264.

Acta Cryst. (1973). B**29**, 1238

The Crystal Structure of Naphtho[*b*]cyclobutene*

BY JAMES L. CRAWFORD† AND RICHARD E. MARSH

Arthur Amos Noyes Laboratory of Chemical Physics, California Institute of Technology, Pasadena, California 91109, *U.S.A.*

(*Received* 28 *December* 1972; *accepted* 5 *February* 1973)

Crystals of naphtho[*b*]cyclobutene, $C_{12}H_{10}$, are monoclinic, space group $P2_1/c$, with $a=5{\cdot}796$ (2), $b=18{\cdot}015$ (8), $c=8{\cdot}291$ (3) Å, $\beta=106{\cdot}50$ (3)°. A structure determination, based on 1228 reflections measured on a diffractometer, led to an R index of 0·035 and estimated deviations of about 0·0015 Å in the coordinates of the C atoms and 0·015 Å in the coordinates of the H atoms. The fusion of the cyclobutene ring results in small distortions in the naphthalene grouping, presumably due to bond-angle constraints at C(2) and C(3).

Introduction

This work is part of a program of studies on the geometries of strained fused-ring compounds.

Experimental

Crystals of naphtho[*b*]cyclobutene, which was first synthesized by Cava & Shirley (1960), were supplied by Dr Ashley Cooper of the Jet Propulsion Laboratory, Pasadena. They were colorless prisms, somewhat elongated along **a**. They sublime slowly at room temperature. The crystal used in this work had dimensions $0{\cdot}36\times0{\cdot}30\times0{\cdot}30$ mm (along **a**, **b** and **c**) and was sealed in a 0·3 mm diameter thin-walled glass capillary to prevent sublimation. Preliminary rotation and Weissenberg photographs showed monoclinic symmetry with systematic absences ($h0l$, l odd; $0k0$, k odd) characteristic of space group $P2_1/c$. The crystal was then mounted on a General Electric quarter-circle diffractometer, automated by Datex and highly modified by Dr Sten Samson. Cell dimensions were obtained from a least-squares fit to 2θ values for 15 reflections; they are given in Table 1. We did not measure the crystal density.

Table 1. *Crystal data*

Naphtho[*b*]cyclobutene	$C_{12}H_{10}$
Monoclinic	F.W. 154·2
Space group, $P2_1/c$	m.p. 86 °C*
$a=5{\cdot}796$ (2) Å	$Z=4$
$b=18{\cdot}015$ (8)	$F(000)=328$
$c=8{\cdot}291$ (3)	$D_x=1{\cdot}23$ g cm^{-3}
$\beta=106{\cdot}50$ (3)°	$\mu=4$ cm^{-1}
$V=830{\cdot}1$ (9) Å^3	λ(Cu $K\alpha$) = 1·5418 Å

* Cava & Shirley (1960).

Intensities were measured using Cu $K\alpha$ radiation and θ–2θ scans at a speed of 2° (in 2θ) per min; backgrounds were counted for 30 sec at the extrema. The $11\bar{2}$ reflection was monitored every 20 reflections; its intensity decreased from 864,000 to 820,000 counts – about 5% – during the one-week period of data collection. All reflections in one quadrant of reciprocal space out to $2\theta=130°$ were surveyed; they numbered 1236, of which 49 had net intensities less than zero. Six reflections ($\bar{5}74$; $\bar{1}$,17,1; $\bar{1}$,17,2; $\bar{1}85$; $\bar{1}95$; and $\bar{1}$,10,5) were not recorded because of failures

* Contribution No. 4599 from the Arthur Amos Noyes Laboratory of Chemical Physics, California Institute of Technology, Pasadena, California 91109.

† Present address: Department of Chemistry, Harvard University, Cambridge, Massachusetts 02138, U.S.A.